# Contents

**Preface**  xi  

1  **Dealing with data**  1  
1.1  The role of statistics in geography  1  
1.1.1  Why do geographers need to use statistics?  1  
1.2  About this book  3  
1.3  Data and measurement error  3  
1.3.1  Types of geographical data: nominal, ordinal, interval, and ratio  3  
1.3.2  Spatial data types  5  
1.3.3  Measurement error, accuracy and precision  6  
1.3.4  Reporting data and uncertainties  7  
1.3.5  Significant figures  9  
1.3.6  Scientific notation (standard form)  10  
1.3.7  Calculations in scientific notation  11  
   Exercises  12  

2  **Collecting and summarizing data**  13  
2.1  Sampling methods  13  
2.1.1  Research design  13  
2.1.2  Random sampling  15  
2.1.3  Systematic sampling  16  
2.1.4  Stratified sampling  17  
2.2  Graphical summaries  17  
2.2.1  Frequency distributions and histograms  17  
2.2.2  Time series plots  21  
2.2.3  Scatter plots  22  
2.3  Summarizing data numerically  24
2.3.1 Measures of central tendency: mean, median and mode 24
2.3.2 Mean 24
2.3.3 Median 25
2.3.4 Mode 25
2.3.5 Measures of dispersion 28
2.3.6 Variance 29
2.3.7 Standard deviation 30
2.3.8 Coefficient of variation 30
2.3.9 Skewness and kurtosis 33
Exercises 33

3 Probability and sampling distributions 37
3.1 Probability 37
3.1.1 Probability, statistics and random variables 37
3.1.2 The properties of the normal distribution 38
3.2 Probability and the normal distribution: z-scores 39
3.3 Sampling distributions and the central limit theorem 43
Exercises 47

4 Estimating parameters with confidence intervals 49
4.1 Confidence intervals on the mean of a normal distribution: the basics 49
4.2 Confidence intervals in practice: the t-distribution 50
4.3 Sample size 53
4.4 Confidence intervals for a proportion 53
Exercises 54

5 Comparing datasets 55
5.1 Hypothesis testing with one sample: general principles 55
5.1.1 Comparing means: one-sample z-test 56
5.1.2 p-values 60
5.1.3 General procedure for hypothesis testing 61
5.2 Comparing means from small samples: one-sample t-test 61
5.3 Comparing proportions for one sample 63
5.4 Comparing two samples 64
  5.4.1 Independent samples 64
  5.4.2 Comparing means: \( t \)-test with unknown population variances assumed equal 64
  5.4.3 Comparing means: \( t \)-test with unknown population variances assumed unequal 68
  5.4.4 \( t \)-test for use with paired samples (paired \( t \)-test) 71
  5.4.5 Comparing variances: \( F \)-test 74
5.5 Non-parametric hypothesis testing 75
  5.5.1 Parametric and non-parametric tests 75
  5.5.2 Mann–whitney \( U \)-test 75
Exercises 79

6 Comparing distributions: the Chi-squared test 81
  6.1 Chi-squared test with one sample 81
  6.2 Chi-squared test for two samples 84
Exercises 87

7 Analysis of variance 89
  7.1 One-way analysis of variance 90
  7.2 Assumptions and diagnostics 99
  7.3 Multiple comparison tests after analysis of variance 101
  7.4 Non-parametric methods in the analysis of variance 105
  7.5 Summary and further applications 106
Exercises 107

8 Correlation 109
  8.1 Correlation analysis 109
  8.2 Pearson’s product-moment correlation coefficient 110
  8.3 Significance tests of correlation coefficient 112
  8.4 Spearman’s rank correlation coefficient 114
  8.5 Correlation and causality 116
Exercises 117

9 Linear regression 121
  9.1 Least-squares linear regression 121
  9.2 Scatter plots 122
9.3 Choosing the line of best fit: the ‘least-squares’ procedure 124
9.4 Analysis of residuals 128
9.5 Assumptions and caveats with regression 130
9.6 Is the regression significant? 131
9.7 Coefficient of determination 135
9.8 Confidence intervals and hypothesis tests concerning regression parameters 137
  9.8.1 Standard error of the regression parameters 137
  9.8.2 Tests on the regression parameters 138
  9.8.3 Confidence intervals on the regression parameters 139
  9.8.4 Confidence interval about the regression line 140
9.9 Reduced major axis regression 140
9.10 Summary 142
Exercises 142

10 Spatial statistics 145
10.1 Spatial data 145
  10.1.1 Types of spatial data 145
  10.1.2 Spatial data structures 146
  10.1.3 Map projections 149
10.2 Summarizing spatial data 157
  10.2.1 Mean centre 157
  10.2.2 Weighted mean centre 157
  10.2.3 Density estimation 158
10.3 Identifying clusters 159
  10.3.1 Quadrat test 159
  10.3.2 Nearest neighbour statistics 162
10.4 Interpolation and plotting contour maps 162
10.5 Spatial relationships 163
  10.5.1 Spatial autocorrelation 163
  10.5.2 Join counts 164
Exercises 171

11 Time series analysis 173
11.1 Time series in geographical research 173
11.2 Analysing time series 174
  11.2.1 Describing time series: definitions 174
  11.2.2 Plotting time series 175